

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Joseph Hamburger and Avraham Laban  
METHODS FOR BULK STABLE  
INTRODUCTION AND EXPRESSION OF  
FOREIGN GENES INTO EUKARYOTIC  
PARASITES FOR IN VIVO PRODUCTION  
OF DESIRED GENE PRODUCTS
- (iii) NUMBER OF SEQUENCES: 9
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
(B) STREET: 2001 Jefferson Davis Highway, Suite 207  
(C) CITY: Arlington  
(D) STATE: Virginia  
(E) COUNTRY: United States of America  
(F) ZIP: 22202
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
(B) COMPUTER: Twinhead® Slimnote-890TX  
(C) OPERATING SYSTEM: MS DOS version 6.2,  
Windows version 3.11  
(D) SOFTWARE: Word for Windows version 2.0
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: 09/201,850  
(B) FILING DATE: December 1, 1998
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Friedmam, Mark M.  
(B) REGISTRATION NUMBER: 33,883  
(C) REFERENCE/DOCKET NUMBER: 325/71
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 972-3-5625553  
(B) TELEFAX: 972-3-5625554  
(C) TELEX:

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 5356  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
- |   |     |
|---|-----|
| AGCTTGCATG CCTGCAGGTC GACTCTAGAG GATCCCCGGG TACCGGTAGA  | 50  |
| AAAAATGAGT AAAGGAGAAG AACTTTTCAC TGGAGTTGTC CCAATTCTTG  | 100 |
| TTGAATTAGA TGGTGATGTT AATGGGCACA AATTTTCTGT CAGTGGAGAG  | 150 |
| GGTGAAGGTG ATGCAACATA CGGAAAACTT ACCCTTAAAT TTATTGCAC   | 200 |
| TACTGGAAAA CTACCTGTTT CATGGCCAAC ACTTGTCAC TTTTCTCTT    | 250 |
| ATGGTGTTCA ATGCTTTTCA AGATACCCAG ATCATATGAA ACGGCATGAC  | 300 |
| TTTTTCAAGA GTGCCATGCC CGAAGGTTAT GTACAGGAAA GAACTATATT  | 350 |
| TTTCAAAGAT GACGGGAAC TACAAGACAC TGCTGAAGTC AAGTTTGAAG   | 400 |
| GTGATACCCT TGTTAATAGA ATCGAGTTAA AAGGTATTGA TTTTAAAGAA  | 450 |
| GATGGAAACA TTCTTGGACA CAAATGGGA TACAATACTATA ACTCACACAA | 500 |
| TGTATACATC ATGGCAGACA AACAAAAGAA TGGAAATCAA GTTAACCTCA  | 550 |
| AAATTAGACA CAACATTGAA GATGGAAGCG TTCAACTAGC AGACCATTAT  | 600 |

CAACAAAATA CTCCAATTGG CGATGGCCCT GTCCTTTTAC CAGACAACCA 650  
 TTACCTGTCC ACACAATCTG CCTTTTCGAA AGATCCCAAC GAAAAGAGAG 700  
 ACCACATGGT CCTTCTTGAG TTTGTAACAG CTGCTGGGAT TACACATGGC 750  
 ATGGATGAAC TATACAAATA GCATTTCGTAG AATTCTATTG CACGGTATAT 800  
 GGCGAAGAAA CATCATATGA TGGGTGAAAC AGACGAGGAA TACTATAGTG 850  
 TTGAAAAGTT GATTGGTCAG GTGAGTTGAG CCTTTTATTG GTTGATGGGA 900  
 TTTTATGCTA CTGGAGTCCT GTAAGTTCAG TTTGTTGCGT TGAATAAATA 950  
 AAGTTCGAAT TGATAGGCGA GTGGTAACCT ATGTGTTGTG ATTACTAAAA 1000  
 GTCATAAGTG TGAGGTATCC CAATTATTAG AATGTGACT AAGACTGAGC 1050  
 TTTTCTGCCT ATTGGGATAT TACAGATAAG TTAATTATTG GTAGTATCGT 1100  
 AATAGTGTTT ACTGGGAACC ATCTAATTCA CTCTCAAAA TACTTTATCG 1150  
 TGTCTTATTT TATTTCATCT CATAACTCGT TTGCTAGGAG AACAGCTGTG 1200  
 CCCATAGGTT AGTTGACAAG TTACTTACTT CCCCATGATA CGCCTTATTT 1250  
 TAATAACCTA ATAAAGGCTC GGTACGTAA CGTCCCTATA GTGCATAGCA 1300  
 TACCTACGAA CATTGACTCC CAACTGATTA AGTCCGGATT GCATTTTTTT 1350  
 GGATGTTATC GCACAGTAAG ACAATACCAT CCTCATCTCA ATGTCAAGAG 1400  
 GTCCTTCTCC AGGCAACAGA TCCATACCAC CATTACTTAC ATCCATCAGA 1450  
 GCTGTTCCCA GAATGTCTTG TTATAAACGG ATGGATTATT TATTTATTTG 1500  
 ACCACATAAA TATTGCATCA AAGAGGTGGG GGATCCACTA GTTCTAGAGC 1550  
 GGCGGCCACC GCGGTGGAGC TCCAGCTTTT GTTCCCTTTA GTGAGGGTTA 1600  
 ATTGCGCGCT TGGCGTAATC ATGGTCATAG CTGTTTCTTG TGTGAAATTG 1650  
 TTATCCGCTC ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA 1700  
 AAGCCTGGGG TGCCTAATGA GTGAGCTAAC TCACATTAAT TGCCTTGCGC 1750  
 TCACTGCCCG CTTCCTCAGTC GGGAAACCTG TCGTGCCAGC TGCATTAATG 1800  
 AATCGGCCAA CGCGCGGGGA GAGCGGTTT GCGTATTGGG CGCTCTTCCG 1850  
 CTTCTCTCGT CACTGACTCG CTGCGCTCGG TCGTTCGGCT GCGGCGAGCG 1900  
 GTATCAGCTC ACTCAAAGGC GGTAAATACG TTATCCACAG AATCAGGGGA 1950  
 TAACGCAGGA AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC 2000  
 GTAAAAAGGC CGCGTTGCTG GCGTTTTTCC ATAGGCTCCG CCCCCTGAC 2050  
 GAGCATCACA AAAATCGACG CTCAAGTCAG AGGTGGCGAA ACCGACAGGA 2100  
 CTATAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC 2150  
 TGTTCGACCT CTGCGCTTA CCGGATACCT GTCCGCTTT CTCCCTTCGG 2200  
 GAAGCGTGGC GCTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTCGGTG 2250  
 TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG CACGACCCCG CGTTCAGCCC 2300  
 GACCGCTGCG CTTTATCCGG TAACTATCGT CTTGAGTCCA ACCCGGTAAG 2350  
 ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG ATTAGCAGAG 2400  
 CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG GCCTAACTAC 2450  
 GGCTACACTA GAAGGACAGT ATTTGGTATC TGCGCTCTGC TGAAGCCAGT 2500  
 TACCTTCGGA AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA CAAACCACCG 2550  
 CTGGTAGCTG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC GCGCAGAAAA 2600  
 AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGT CTGACGCTCA 2650  
 GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCAATGAGA TTATCAAAAA 2700  
 GGATCTTCAC CTAGATCCTT TTAATTAATA AATGAAGTTT TAAATCAATC 2750  
 TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT GCTTAATCAG 2800  
 TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTTATCC ATAGTTGCCT 2850  
 GACTCCCCGT CGTGTAGATA CTACGATACG GGAGGGCTTA CCATCTGGCC 2900  
 CCAGTGCTGC AATGATACCG CGAGACCCAC GCTCACCAGC TCCAGATTTA 2950  
 TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCTGC 3000  
 AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG 3050  
 TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTACA 3100  
 GGCATCGTGG TGTACGCTC GTCGTTGGT ATGGCTTCAT TCAGTCCGG 3150  
 TTCCCAACGA TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAAG 3200  
 CGGTTAGCTC CTTGCTCCT CCGATCGTTG TCAGAAAGTAA GTTGGCCGCA 3250  
 GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC TTAAGTTCAT 3300  
 GCCATCCGTA AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT 3350  
 TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC GCGTCAATA 3400  
 CGGGATAATA CCGCGCCACA TAGCAGAACT TAAAAAGTGC TCATCATTGG 3450  
 AAAACGTTCT TCGGGGCGAA AACTCTCAAG GATCTTACCG CTGTTGAGAT 3500

```

CCAGTTCGAT GTAACCCACT CGTGACCCCA ACTGATCTTC AGCATCTTTT 3550
ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC 3600
AAAAAAGGGA ATAAGGGCGA CACGGAATG TTGAATACTC ATACTCTTCC 3650
TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA 3700
TACATATTTG AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC 3750
ATTTCCTCGA AAAGTGCCAC CTGACGCGCC CTGTAGCGGC GCATTAAGCG 3800
CGGCGGGTGT GGTGGTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC 3850
CTAGCGCCCG CTCCTTTTCG TTTCTTCCCT TCCTTTCTCG CCACGTTTCG 3900
CGGCTTTCCC CGTCAAGCTC TAAATCGGGG GCTCCCTTTA GGGTTCCGAT 3950
TTAGTGCTTT ACGGCACCTC GACCCCAAAA AACTTGATTA GGGTGATGGT 4000
TCACGTAGTG GGCCATCGCC CTGATAGACG GTTTTTCGCC CTTTGACGTT 4050
GGAGTCCACG TTCTTTAATA GTGGACTCTT GTTCCAACT GGAACAACAC 4100
TCAACCTTAT CTCGGTCTAT TCTTTTGATT TATAAGGGAT TTTGCCGATT 4150
TCGGCCTATT GGTAAAAAAA TGAGCTGATT TAACAAAAAT TTAACGCGAA 4200
TTTAACAAA ATATTAAACG TTACAATTTT CATTCGCCAT TCAGGCTGCG 4250
CAACTGTTGG GAAGGGCGAT CGGTGCGGGC CTCTTCGCTA TTACGCCAGC 4300
TGGCGAAAGG GGGATGTGCT GCAAGGCGAT TAAGTTGGGT AACGCCAGGG 4350
TTTTCCAGT CACGACGTTG TAAAACGACG GCCAGTGAGC GCGCGTAATA 4400
CGACTCACTA TAGGGCGAAT TGGGTACCGG GCGGCTCCTC GAGGTCGACG 4450
GTATCGATAA GCTTGATTCA TCGAGAACGG TTTACATGTT CAATGAATCG 4500
AGTCAAATTT GTCTGCTTAA TTTTATTGG TCACTCTTTC ACAGCCAATG 4550
AGGCACTCAA TAAACAGCGA ATAGAAATGA AATATTTACA GTTAAAAATCA 4600
AGAGTTACAC TATGGCCGA TCTGTTTACT AAATGACTAT TTAATAGTCA 4650
CAGCTCAGAT AGTTAGCACT GTTTTGTCTA TTTGCAAGAT GGCTGGCGAG 4700
CATATCAAGG TGGGTTGGTG TGGTGTTCG TCATTGTGTT GCTTGGTAAA 4750
TTTTCGAGTG TTGTTAAGGT TCCTAATTGC AGCATTAGC GGAATAAATA 4800
TTATTCGCT ATATGAAGGT AAATACTGAG TCCATGTAA TTGCACATTA 4850
GATCCGTACA TATATTGGC TGTGATTTTC CACGTCATCA GGTGAGTAA 4900
TTACGTGTTT AACTGCGGTT CTGATATAAT GCACAACCAT TATAGTCACA 4950
TAGTGGTCAC CGATTATGTG TTAGGTGTAG CATTTCAGGC GATGTTATAG 5000
TTTGTCTGTG GAATTGTTAC GAGAACTTAG GGGATTCTAG AGCGTGAATC 5050
CATCTAGTCC ATATCTCAGA GACATGCGTC CATGGAACAG TTATTTATTA 5100
TGTGATGTTT TGTCTGAAGT GGTTTTTTTC AAGCGTAAAT TGATGTGAAA 5150
CACAGCTGTA ATTTCTATAG GTTATCTATT TTGACGGACG CGGACGTGCT 5200
GAATCGATTG GGATGACTCT TGTGGCAGCT GGTGTAGACT ACGAAGATGA 5250
GAGAAATAGT TTCCAAGATT GGCCAAAAAT CAACCAACT ATTCCAGGCG 5300
GACGATTGCC TGCAGTGAAG GTCAGTATG ATCATGGGCA CGTGAAATGG 5350
ATGTTA 5356

```

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6494  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

GATCTGAATC CGACCAACCG TTCTATGAAA ATCGTTGTAT CTCGGAACCC 50
ACTGGACGGA TTTTATGAT GTTTGTTTTA GATTATTTGC GAGAGCGTGG 100
GCGTTAATAT AAAACAAGAA TGATCTGAAT CCGACCAACC GTTCTATGAA 150
AATCGTTGTA TCTCCGAAAC CACTGGACGG ATTTTATGA TGTGTTTTT 200
AGATTATTTG CGAGAGCGTG GCGTTAATA TAAAACAAGA ATGATCTGAA 250
TCCGACCAAC CGTTCTATGA AAATCGTTGT ATCTCCGAAA CCACTGGACG 300
GATTTTTATG ATGTTTGTTC TAGATTATTT GCGAGAGCGT GGGCGTTAAT 350
ATAAAACAAG AATCATCTCA ATCCATCAG CCGTTCTATG AAAATCGTTG 400
TATCTCCGAA ACCACTGGAC GGATTTTTAT GATGTTTGTT TTAGATTATT 450
TGCGAGAGCG TGGGCGTTAA TATAAAACAA GAATGATCTG AACACGGGTG 500
TTTTTCTGTT CAGCTTATGC AACTTTAAAA TTCGATGGGT CGTCTCAACG 550
AAATTGTGAT TGCTTTGTCG AGGTCGACGG TATCGATAAG CTTGATTCAT 600
CGAGAACGGT TTACATGTTT AATGAATCGA GTCAATTTG TCTGCTTAAT 650
TTTTATTGGT CACTCTTTCA CAGCCAATGA GGCCTCAAT AAACAGCGAA 700

```

TAGAAATGAA	ATATTTACAG	TTAAAATCAA	GAGTTACACT	ATTGGCCGAT	750
CTGTTTACTA	AATGACTATT	TAATAGTCAC	AGCTCAGATA	GTTAGCACTG	800
TTTTGTCTAT	TTGCAAGATG	GCTGGCGAGC	ATATCAAGGT	GGGTTGGTGT	850
GGTTGTTTCG	CATTGTGTTG	CTTGGTAAAT	TTTCGAGTGT	TGTTAAGGTT	900
CCTAATTGCA	GCATTTAGCG	GAATAAATAT	TATTCGCTA	TATGAAGGTA	950
AATACTGAGT	CCATGTTAAT	TGCACATTAG	ATCCGTACAT	ATATTTGGCT	1000
GTGATTTTCC	ACGTCATCAG	GTCAGTAAGT	TACGTGTTCA	CACTGGGTTC	1050
TGATATAATG	CACAACCATT	ATAGTCACAT	AGTGGTCACC	GATTATGTGT	1100
TAGGTGTAGC	ATTTCAGGCG	ATGTTATAGT	TTGTCTGTGG	AATTGTTACG	1150
AGAACTTAGG	GGATTCTAGA	GCGTGAATCC	ATCTAGTCCA	TATCTCAGAG	1200
ACATGCGTCC	ATGGAACAGT	TATTTATTAT	GTGATGTTTT	GTCTGAAGTG	1250
GTTTTTTTCA	AGCGTAAATT	GATGTGAAAC	ACAGCTGTAA	TTTCTATAGG	1300
TTATCTATTT	TGACGGACGC	GGACGTGCTG	AATCGATTCT	GATGACTCTT	1350
GTGGCAGCTG	GTGTAGACTA	CGAAGATGAG	AGAATTAGTT	TCCAAGATTG	1400
GCCAAAATC	AAACCAACTA	TTCCAGGCGG	ACGATTGCCT	GCAGTGAAAG	1450
TCCTGATGA	TCATGGGCAC	GTGAAATGGA	TGTTAAGCTT	GCATGCCTGC	1500
AGGTGCACTC	TAGAGGATCC	CCGGGTACCG	GTAGAAAAAA	TGAGTAAAGG	1550
AGAAGAACTT	TTCACTGGAG	TTGTCCCAAT	TCTGTGTGAA	TTAGATGGTG	1600
ATGTTAATGG	GCACAAATTT	TCTGTCAGTG	GAGAGGGTGA	AGGTGATGCA	1650
ACATACGGAA	AACCTTACCCT	TAAATTTATT	TGCACTACTG	GAAAACTACC	1700
TGTTCCATGG	CCAACACTTG	TCCTACTTTT	CTCTTATGGT	GTTCAATGCT	1750
TTTCAAGATA	CCAGATCAT	ATGAAACGGC	ATGACTTTTT	CAAGAGTGCC	1800
ATGCCCGAAG	GTTATGTACA	GGAAAGAAGT	ATATTTTTC	AAGATGACGG	1850
GAACCTACAAG	ACACGTGCTG	AAGTCAAGTT	TGAAGGTGAT	ACCCTTGTTA	1900
ATAGAATCGA	GTTAAAAGGT	ATTGATTTTA	AAGAAGATGG	AAACATTCTT	1950
GGACACAAAT	TGGAATACAA	CTATAACTCA	CACAATGTAT	ACATCATGGC	2000
AGACAAACAA	AAGAATGGAA	TCAAAGTTAA	CTTCAAAAT	AGACACAACA	2050
TTGAAGATGG	AAGCGTTCAA	CTAGCAGACC	ATTATCAACA	AAATACTCCA	2100
ATTGGCGATG	GCCCTGTCCCT	TTTACCAGAC	AACCATTACC	TGTCCACACA	2150
ATCTGCCCTT	TCGAAAGATC	CCAACGAAAA	GAGAGACCAC	ATGGTCCTTC	2200
TTGAGTTTGT	AACAGCTGCT	GGGATTACAC	ATGGCATGGA	TGAACTATAC	2250
AAATAGCATT	CGTAGAATTC	TATTGCACGG	TATATGGCGA	AGAAACATCA	2300
TATGATGGGT	GAAACAGACG	AGGAATACTA	TAGTGTGAA	AAGTTGATTG	2350
GTCAGGTGAG	TTGAGCCTTT	TATTGGTTGA	TGGGATTTTA	TGCTACTGGA	2400
GTCCTGTAAG	TTCAGTTTGT	TGCGTTGAAT	AAATAAAGTT	CGAATTGATA	2450
GGCGAGTGGT	AACCTATGTG	TTGTGATTAC	TAAAGTCAAT	AAGTGTGAGG	2500
TATCCCAATT	ATTAGAAATG	TGACTAAGAC	TGAGCTTTTC	TGCCTATTGG	2550
GATATTACAG	ATAAGTTACT	TATTTGTAGT	ATCGTAATAG	TGTTCACTGG	2600
GAACCATCTA	ATTCACCTCT	AAAAATACTT	TATCGTGTCT	TATTTTATTTC	2650
ATCCTCATAA	CTCGTTTGCT	AGGAGAACAG	CTGTGCCCAT	AGGTAGATTG	2700
ACAAGTTACT	TACTTCCCCA	TGATACGCCT	TATTTTAATA	ACCTAATAAA	2750
GGCTCGGTTA	CGTAACGTCC	CTATAGTGCA	TAGCATACCT	ACGAACATTG	2800
ACTCCCAACT	GATTAAGTGC	GGATTGCATT	TTTTTGGATG	TTATCGCACA	2850
GTAAGACAAT	ACCATCCTCA	TCTCAATGTC	AAGAGGTCTT	TCTCCAGGCA	2900
ACAGATCCAT	ACCACCATTA	CTTACATCCA	TCAGAGCTGT	TCCCAGAATG	2950
TCTTGTTATA	AACGGATGGA	TTATTTATTT	ATTTGACCAC	ATAAATATTG	3000
CATCAAAGAG	GTGGGGGATC	CACTAGTTCT	AGAGCGGCCG	CCACCGCGGT	3050
GGAGGATCTG	AATCCGACCA	ACCGTTCTAT	GAAAATCGTT	GTATCTCCGA	3100
AACCACTGGA	CGGATTTTTA	TGATGTTTGT	TTTAGATTAT	TTGCGAGAGC	3150
GTGGGCGTTA	ATATAAAACA	AGAATGATCT	GAATCCGACC	AACCGTTCTA	3200
TGAAATCGT	TGTATCTCCG	AAACCACTGG	ACGGATTTT	ATGATGTTTG	3250
TTTTAGATTA	TTTGGCGAGG	CGTGGCGGTT	AATATAAAAC	AAGAATGATC	3300
TGAATCCGAC	CAACCGTTCT	ATGAAAATCG	TTGTATCTCC	GAAACCACTG	3350
GACGGATTTT	TATGATGTTT	GTTTTAGATT	ATTTGCGAGA	GCGTGGGCGT	3400
TAATATAAAA	CAAGAATCAT	CTCAATCCCA	TCAGCCGTTT	TATGAAAATC	3450
GTTGTATCTC	CGAAACCACT	GGACGGATTT	TTATGATGTT	TGTTTTAGAT	3500
TATTTGCGAG	AGCGTGGCG	TTAATATAAA	ACAAGAATGA	TCTGAACACG	3550
GGTGTTTTTC	TGTTTCAGCTT	ATGCAACTTT	AAAATTCGAT	GGGTCTCTCT	3600
AACGAAATTT	GTAATTGCTTT	GCTCCAGCTT	TTGTTCCCTT	TAGTGAGGGT	3650
TAATTGCGCG	CTTGGCGTAA	TCATGGTCAT	AGCTGTTTCC	TGTGTGAAAT	3700
TGTTATCCGC	TCACAAATCC	ACACAACATA	CGAGCCGGAA	GCATAAAGTG	3750
TAAAGCCTGG	GGTGCCCTAAT	GAGTGAGCTA	ACTCACATTA	ATTGCGTTGC	3800
GCTCACTGCC	CGCTTTCCAG	TCGGGAAACC	TGTCGTGCCA	GCTGCATTAA	3850
TGAATCGGCC	AACGCGCGGG	GAGAGCGGTT	TTGCGTATTG	GGCGCTCTTC	3900
CGCTTCTCTG	CTCACTGACT	CGCTGCGCTC	GGTCGTTTCG	CTGCGGCGAG	3950
CGGTATCAGC	TCACTCAAAAG	GCGGTAATAC	GGTTATCCAC	AGAATCAGGG	4000
GATAACGCG	GAAAGAACAT	GTGAGCAAAA	GGCCAGCAAA	AGGCCAGGAA	4050
CGGTAAAAAG	GCCGCGTTGC	TGGCGTTTTT	CCATAGGCTC	CGCCCCCTTG	4100
ACGAGCATCA	CAAAAATCGA	CGCTCAAGTC	AGAGGTGGCG	AAACCGACAG	4150
GACTATAAAG	ATACGAGGCG	TTTCCCCCTG	GAAGCTCCCT	CGTGCCTCT	4200

```

CCTGTTCCGA CCCTGCCGCT TACCGGATAC CTGTCCGCCT TTCTCCCTTC 4250
GGGAAGCGTG GCGCTTTCTC ATAGCTCACG CTGTAGGTAT CTCAGTTCGG 4300
TGTAGGTCGT TCGCTCCAAG CTGGGCTGTG TGCACGACCC CCCGTTACGC 4350
CCGACCGCTG CGCCTTATCC GGTAACATC GTCTTGAGTC CAACCCGGTA 4400
AGACACGACT TATCGCCACT GGCAGCAGCC ACTGGTAACA GGATTAGCAG 4450
AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG TGGCCTAACT 4500
ACGGCTACAC TAGAAGGACA GTATTGGTA TCTGCGCTCT GCTGAAGCCA 4550
GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC 4600
CGCTGGTAGC GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA 4650
AAAAAGGATC TCAAGAAGAT CCTTTGATCT TTTCTACGGG GTCTGACGCT 4700
CAGTGGAACG AAAAATCACG TTAAGGGATT TTGGTCATGA GATTATCAAA 4750
AAGGATCTTC ACCTAGATCC TTTTAAATTA AAAATGAAGT TTTAAATCAA 4800
TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC 4850
AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTTCT CATAGTTGTC 4900
CTGACTCCCC GTCGTGTAGA TACTACGATA CGGGAGGGCT TACCATCTGG 4950
CCCCAGTGTG GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT 5000
TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT 5050
GCAACTTTAT CCGCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG 5100
AGTAAGTAGT TCGCCAGTTA ATAGTTTGGC CAACGTTGTT GCCATTGCTA 5150
CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC 5200
GGTTCCTAAC GATCAAGGCG AGTTACATGA TCCCCCATGT TGTGCAAAAA 5250
AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAACT AAGTTGGCCG 5300
CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC 5350
ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC 5400
ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA 5450
TACGGGATAA TACCGGCCCA CATAGCAGAA CTTTAAAGT GCTCATCATT 5500
GGAAACGTT CTTCGGGGCG AAAATCTCTA AGGATCTTAC CGCTGTTGAG 5550
ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACGATCT TCAGCATCTT 5600
TTACTTTTAC CAGCGTTTCT GGTGAGCAA AAACAGGAAG GCAAAATGCC 5650
GCAAAAAAGG GAATAAGGGC GACACGAAA TGTTGAATAC TCATACTCTT 5700
CCTTTTTCAT TATTATTGAA GCATTATCA GGGTTATTGT CTCATGAGCG 5750
GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCGCGC 5800
ACATTTCCCC GAAAAGTGCC ACCTGACGCG CCCTGTAGCG GCGCATTAG 5850
CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG 5900
CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC CTTCTTTTCT CGCCACGTTT 5950
GCCGCTTTC CCCGTCAAGC TCTAAATCGG GGGCTCCCTT TAGGGTCCG 6000
ATTTAGTGCT TTACGCGACC TCGACCCCAA AAAACTTGAT TAGGGTGATG 6050
GTTACAGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG CCCTTTGACG 6100
TTGGAGTCCA CGTTCTTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC 6150
ACTCAACCT ATCTCGGTCT ATTCTTTTGA TTTATAAGGG ATTTTGCCGA 6200
TTTCGGCCTA TTGGTTAAAA AATGAGCTGA TTTAAACAAA ATTTAACGCG 6250
AATTTTAAAC AAATATTAAC GCTTACAATT TCCATTGCGC ATTCAGGCTG 6300
CGCAACTGTT GGGAAAGGCG ATCGGTGCGG GCCTCTTCGC TATTACGCCA 6350
GCTGGCGAAA GGGGGATGTG CTGCAAGGCG ATTAAGTTGG GTAACGCCAG 6400
GGTTTTCCCA GTCACGACGT TGTAACGA CGGCCAGTGA GCGCGCGTAA 6450
TACGACTCAC TATAGGCGA ATTGGGTACC GGGCCCCCCC TCGA 6494

```

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CATCGAGAAC GGTTTACATG TTCA 24

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCAGCTCTCT CACATGCTCC A 21

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATGAGTAAAG GAGAAGAACT TTTC 24

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTTGTATAGT TCATCCATGC C 21

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TCTCCCATGA TGTATACATT ATGT 24

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TCTCCATCGA AGGGTCATCA CG 22

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 121  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GATCTCACCT	ATCAGACGAA	ACAAAGAAAA	TTTTAAAT	GTTGGTGGAA	50
GTGCCTGTTT	CGCAATATCT	CCGGAATGGT	TGGTCGTATC	GTTGTGAAAA	100
TTGTTTCATA	TTATTGGTGA	C			121